Evaluation of Reference Marker Geometries for LiDAR Pose Estimation

Richard B. Morrison & Mathias Kölsch
Naval Postgraduate School
Monterey, California
rbmorris@nps.edu

ABSTRACT

Light Detection and Ranging (LiDAR) devices are increasingly used for sensing during training and operations. One difficulty is the estimation of its exact pose (position and orientation), particularly in the absence of good GPS position data and inertial measurement unit (IMU) orientation data. Placing reference markers (fiducials) at known positions permits the use of the LiDAR range measurement data *alone* to derive relative pose estimation, *without* additional sensors.

The purpose of this study is to determine the suitability of various 3D geometries for their use as reference markers. Specifically, these geometries reviewed for the desirable qualities of LiDAR detection at various distances and orientations and for the potential to determine the precise relative position and orientation of the marker and LiDAR. The intensity of laser range returns are not considered in this study. In related work, 2D high contrast markers have been used to register augmented reality imagery with the environment. Vertical cylinders have been used for pose estimation of a ground-based LiDAR used for an autonomous wheelchair docking system. Three-dimensional spherical markers have provided ground truth data for registering patient movement during medical procedures. However, a thorough, theoretical and practical evaluation of marker geometries has not been performed for ground vehicle LiDAR systems. Through an evaluation of the recognition of the shape and orientation of various 3D geometries using LiDAR range data, we show which geometries overcome the aspect-variant limitations of 2D markers, the ambiguity of vertical cylinders to height and the limitations of spheres to distance scalability. We propose to register camera video with LiDAR range data for quick creation of textured 3D models for use in a ground-vehicle Augmented Reality system.

ABOUT THE AUTHOR

Richard B. Morrison is a 1995 graduate of Worcester Polytechnic Institute with a B.S in Mechanical Engineering and a Lieutenant Commander in the U.S Navy. A Naval Aviator flying the SH-60B Seahawk helicopter, Lieutenant Commander Morrison is qualified as a Seahawk Weapons and Tactics Instructor and has deployed to the Western Pacific, Mediterranean Sea, and Afghanistan. He is currently a Master's degree student in the Modeling, Virtual Environments, and Simulation (MOVES) program at the Naval Post Graduate School in Monterey, CA.

Mathias Kölsch (Ph.D., University of California, Santa Barbara, 2004) is an Assistant Professor of Computer Science at the NPS in Monterey, CA. He is also affiliated with the MOVES Institute, Chair of the MOVES Academic Committee and the Academic Associate for the MOVES curriculum. His research interests include computer vision, hand gesture recognition, augmented and virtual environments, sensor networks and mobile/embedded computing.

Evaluation of Reference Marker Geometries for LiDAR Pose Estimation

Richard B. Morrison & Mathias Kölsch
Naval Postgraduate School
Monterey, California
rbmorris@nps.edu

INTRODUCTION

Range data is becoming more commonplace with sensors such as LiDAR (light detection and ranging), sonar, and even with good stereoscopic imaging sensors and algorithms. Such sensors are more powerful in the type of information they deliver and a surge in their use is expected in the medium timeframe. Their employment is likely to mimic that of electro-optical sensors (video cameras), both with regards to market population and application breath. One category of applications that range sensorsare particularly well suited for is navigation and mapping. For these applications, proper registration across time and space are key capabilities. One approach is the acquisition of good metadata from sources such as Global Positioning System (GPS) and Inertia Measurement Units (IMU), which were particularly important for EO sensors. LiDAR and other sources of range data have far superior precision to GPS (millimeters as opposed to decimeters), yet the accuracy is not as good. The logical choice is to use the sensor to determine its own position relative over time and space. This can be achieved by detecting and tracking particularly prominent features in the environment (landmarks). The drawback of this method is the reliance on the availability of suitable landmarks. Another approach is to manually place fiducial markers into the environment in locations that provide good detection and precise localization.

In this paper, we present the state of the art of fiducials, of methods for their detection, their localization, and their discrimination. The advantages and drawbacks of these methods are described, giving the reader an introduction to the topic, its difficulties, and ways to circumvent them.

The potential of LiDAR enhanced augmented reality (AR) for training applications is the ability to interact with live video that includes AR information tags that are registered in 3D to the video image. An appropriately defined 3D marker geometry could be used as a well-defined control point to register LiDAR sensor data with the video data. Such a 3D marker geometry would need to have characteristics that make it identifiable in both the 3D LiDAR range data as well as the 2D video image. A 3D reference geometry which can be accurately detected with position and orientation could be used for LiDAR pose estimation.

Two Dimensional Augmented Reality Registration

Vision based AR tracking and registration has been accomplished using 2D planar markers with high contrast color patterns that are detected using computer vision algorithms. [Kato, 1999] and [Owen, 2009] Examples of 2D fiducials that have been used include LED lights, multi-colored concentric circles, and two-dimensional bar codes or matrix codes. These 2D markers provide a degree of usability with computer vision detection algorithms, but are susceptible to lighting, scale, and foreshortening variations making them difficult to use for tracking in an outdoor application. [Steinbis, 2008]. To use a LiDAR sensor with a 2D marker, laser pulse intensity measurements would need to be analyzed to detect the marker pattern, which is not always practical in an unknown outdoor environment.

LiDAR Systems

A LiDAR system produces a large amount of data without context. The range data obtained from the LiDAR system contains range and intensity values. This 1D range data is obtained over a 2 dimensional scan pattern to give a 3D representation of surfaces in the environment. It is more difficult to segment 3D range data to detect features than for 2D images. Real-time processing of 3D data is even more difficult due to the wide variety of LiDAR systems available with different range acquisition formats. This data tends to be noisy giving difficulty to less robust fitting algorithms.

LiDAR systems vary in their method of range data acquisition. Flash LiDAR uses a broad field of illumination with a single detector that acquires data simultaneously. A scanning LiDAR uses a single laser with an articulated mirror or an array of lasers that are moved to produce a scan pattern. LiDAR systems used for aerial surveillance of the ground use a sweeping motion while other systems used on ground vehicles use a rotating motion. The type of LiDAR system used affects the format of the measured range data.

For real-time outdoor applications, a LiDAR system must have a high frame rate, large scanning field of regard, and accurate range and angular resolutions. Position information is commonly provided with a combination of GPS and IMU data.

RELATED WORK

Prior work related to pose estimation using 3D geometric markers includes an autonomous docking wheelchair system described in [Gao 2008] and [Gao 2009], medical imaging registration using marker pins with spherical tips [Zhao 1996], and a visual tracking system using cones to estimate the pose of a handheld AR display [Steinbis 2008].

The autonomous docking wheelchair uses two vertical cylinders as markers on the wheelchair and a 2D LiDAR on the docking platform. The LiDAR detects the bearing and range of the cylinders to determine the orientation on the ground plane of the wheelchair.

The medical imaging marker pins have spherical ends that are fixed to the patient's body to track patient movement during surgical operations. This system relies on constraining the motion of the patient and the sensor.

In order to provide an AR display overlay of spinning propelors on a static C-130, computer vision techniques are used to track the orientation of 3D conical shapes placed around the parked aircraft.

LiDAR range mapping is used for autonomous robot simultaneous localization and mapping (SLAM). Papers discussing geometric feature fitting of range data include [Arras, 1997], [Lalonde, 2006], [Nguyen, 2005], and [Gachter, 2005].

A future application for 3D fiducials could be to register 3D LiDAR range data with 2D video image data or stereoscopic image data in an augmented reality system. By segmenting the range data according to range discontinuities and applying various feature extraction methods, the 3D range data points can be translated into 3D polygonal representations. Feature extraction gives the ability to give a description of the object based on the available range data offering a means to compress the range data. [Premebida, 2005] A compressed format, such as a polygonal representation, may be more suitable for use by an augmented reality application.

Comparison of 3D Fiducials

This section compares 3D geometries for suitability as fiducial markers in registration of 3D LiDAR range data with other types of 2D image data, these 3D fiducial shapes are compared for their complexity, robustness of detection algorithms, and applicability to both 2D and 3D sensor domains. A summary is given in Table 1.

Two-Dimensional Markers

2D markers are limited in application to 3D LiDAR range data in that there is no ability to differentiate the 2D marker from any other flat surface in the environment.

Sphere

A sphere offers a symmetrical cross section in all directions and orientation, making pose estimation more difficult. Multiple spheres would need to be used and simultaneously detected to provide enough information for pose estimation. Spheres do offer the advantage of a less complex set of parameters for data point fitting than other geometries. Spheres have a constant circular cross section that can be used for detection algorithms in 2D images provided there is enough contrast with the background.

Cone

A vertically oriented cone offers a symmetrical 360 degree return cross section as well as an identifiable cross section and edge orientation that can be used to identify the position of the cone in 2D images. The intersection of the LiDAR laser scan line with the cone forms a conic section. Feature extraction algorithms can use this fact to identify data points that are potentially associated with the surface of a cone.

Table 1. 3D Geometric Marker Attributes, Advantages, and Disadvantages

Geometry	Complexity	Advantages	Disadvantages
Sphere	Complexity = 4, [radius, center (x,y,z)]	• Low complexity • Circular cross section for detection in 2D • Circle fitting algorithms used for detection	• Ambiguous orientation information. • Multiple markers required for pose estimation.
Cylinder	Complexity = 8 [radius, center of base(x,y,z), height, axis orientation(x,y,z)]	• Quadrilateral cross section for detection in 2D • Possible to use circle finding algorithms for detection if axis is perpendicular to scan line.	• Pose ambiguity along cylinder axis • Possible conic LiDAR scan line return if axis is not perpendicular to scan line.
Cone	Complexity = 8 [radius, center of base (x,y,z), height, orientation of axis(x,y,z)]	• Triangular cross section for detection in 2D • No pose ambiguity along axis. • Possible to use circle finding algorithms for detection if axis is perpendicular to scan line.	• Possible conic LiDAR scan line return if axis is not perpendicular to scan line.
2D planar surface	Complexity = 6 Point on surface (x,y,z) and normal vector (x,y,z)	• Possible for line segment grouping.	

Cylinder

Cylinders have a 360 degree symmetry about an axis offering another conic intersection with the LiDAR laser scan line. Orientation of the axis can assist in pose estimation. Cylinders have the disadvantage of ambiguity of the distance from the base to particular cross section. For example, given a cylindrical cross section along a long pole, without an estimated location for one of the pole ends, the position of the sensor relative to that pole cannot be estimated.

Corners and Edges

In urban or other man-made environments, it could be possible to detect structure corners or edges to assist in estimating pose information. Corners can be identified by the intersection of planar surfaces and an edge can be identified by a large range discontinuity between successive range points. Corner or edge features are also candidates for detection within 2D image data. For registration, the issue becomes determining the degree of certainty with which a corner located in the 3D data matches with a corner located in the 2D data.

Segmentation

Range data scan line segmentation can be accomplished by measuring successive data point ranges for discontinuities. Segmentation identifies and separates range data points that are associated with objects in the environment. With segmented successive data points, feature extraction algorithms can be made more efficient by limiting the range data subsets used. [Premebida, 2005]

Primitive Feature Extraction Algorithms

A major problem to overcome in detecting a 3D geometry in LiDAR range data is that there is no efficient direct 3D geometry fitting for a real-time application. The choice of algorithm also depends on the format and sequence of the data provided by the LiDAR system.

Feature extraction of geometric primitives from raw range point data gives a description of the detected object based on the available range data. [Premebida, 2005] Computer vision algorithms for feature extraction exist mainly for 2D rasterized data. Constraints are often used to limit the orientation of the image or the complexity of the background to assist in processing. Typical feature extraction algorithms depend on segmenting data points, often by detecting edges between dissimilar data values such as color or intensity values. For LiDAR range data, this segmentation can be accomplished through the detection of range discontinuities. [Adams, 1999] With the range data segmented into locally similar values, further analysis of the data composition for feature extraction can be more efficient absent the outlying data.

Once primitive features are detected in the raw scan line range data, higher order processing can be used for extraction of 3D features. For example, lines can be fit to common polygonal surfaces and circular arc segments can be grouped as part of a potential sphere, cylinder, or cone.

Line Primitive Fitting

A common method to find linear edges in 2D intensity images is the Hough transform (HT). This method requires global knowledge of the data making it impractical for the large number of range values obtained by LiDAR systems. The method fits lines to the initial set of data adding instances of fit lines to an accumulator array. The Hough transform also does not account for noise and uncertainty. [Nguyen, 2007]

The incremental algorithm (IA) is another common method used in 2D image line fitting. This method relies on local information and is very simple to implement, however, results have shown it to be less effective than other methods. Other methods such as Split-and-Merge (SM) and Iterative End Point Fit (IEPF) could also be implemented easily with potentially better results, but at the cost of more processing time. [Nguyen, 2007] Recursive Line Fitting (RLF) is another modification of the Split-and-Merge algorithm given in [Gachter, 2005].

Circle Primitive Fitting

In the case of finding circle primitives, the fitting of noisy data tends to result in a large variance as the arc angle approaches zero. The Hough transform method exhibits a weakness with noisy data in that there needs to be enough correct data in the subset for the algorithm to work. Orthogonal Distance Regression minimizes the sum of square distances from each data point to the point on the arc. This method is also sensitive to outliers and can be computationally difficult. Geometric methods can calculate a circluar arc definition with three non-colinear points that are on the circle. However, the resulting matrix can be singular if the points are part of too small an arc. The accuracy of the Kasa method is related to the noise present in the data and the arc length traversed by the data set. Using the Kasa method, arcs with larger radii fit better than with smaller radii. [Rusu, 2003]

The Taubin method for fitting two-dimensional data to a circle has been shown to be robust and work well over a small arc distance of the circle. [Taubin, 1991] [Fitzgibbon, 1999] and [Faber, 2001]. The implementation of this method is also readily available.

Elliptical Section Primitive Fitting

For cylindrical and cone geometries, a laser scan line of range data may return a conic section such as an ellipse if the geometry axis is not perpendicular to the LiDAR scan line. The Taubin method is also applicable to fitting ellipses [Taubin, 1991].

In [Fitzgibbon, 1999] a direct least-squares based approach is presented for fitting ellipses. This method uses segmented data and fits the ellipse parameters using a generalized eigenvector approach. The main disadvantage for this method is that the computation is unstable and has the potential to not return a fit, even when the data represents an ellipse.

Plane Fitting

An algorithm for the extraction of planar regions by analyzing the previously fit line segments is given in [Gachter, 2005]. Called Scan Line Grouping, this algorithm extends line extraction by testing neighboring line segments for fit to a plane. A planar region is then grown by adding subsequent neighboring line segments that also fit the planar region.

Range data could also be fit to a plane with a least squares fitting algorithm, but the difficulty becomes the identification of a segmented region that has a high probability of fit. Without this determination, a great deal of processing would be wasted to calculate all potential planes in the data.

Classification of Primitives

Classification of primitive features has been proposed such that subsets of LiDAR scan line data are labeled into 3 classes: "scatter" representing porous volumes such as trees and vegetation; "linear" articles such as wires and tree branches; and "surfaces" representing solid objects. [Lalonde, 2006] Another Classification method is presented in [Arras, 1997] that is based on gaps, hidden corners, openings, apertures, concave corners, and convex corners assuming a predominantly structured or man-made environment.

CONCLUSIONS

The purpose of this paper was to investigate which 3D marker geometries are best able to allow LiDAR pose estimation and registration of LiDAR range data with 2D video images. Making use of the LiDAR range information to perform accurate registration gives the potential for augmentation of camera image data with LiDAR range data.

Upon a review of previous related work, 3D cone geometries appear to offer the best overall characteristics for registration of range data to image data. Cylindrical markers are limited by ambiguity along their axis and the location of a spherical marker does not give enough information to estimate LiDAR pose.

Further experimentation with LiDAR 3D geometric detection algorithms should be completed to determine the feasibility for use in registering 3D range data to 2D image data. The experimentation would also reveal appropriate LiDAR output data formats that avoids processing constraints associated with excessive amounts of data.

ACKNOWLEDGEMENTS

The authors would like to thank the NPS Remote Sensing Center for use of their Velodyne HDL-64E S2 LiDAR system. Also, we would like to thank MAJ Brett Lindberg for his discussion and insights into potential military applications for ground vehicle augmented reality systems.

REFERENCES

Adams, M.D. (1999). *Sensor Modelling and Design and Data Processing for Autonomous Navigation. World Scientific Series in Robotics and Intelligent Systems - Vol. 13.* New Jersey: World Scientific Publishing.

Arras, K.O., & Sieqwart, R.Y. (1997). Feature Extraction and Scene Interpretation for Map-Based Navigation and Map Building. *Proceedings of SPIE, Mobile Robotics XII, Vol. 3210.*

Cheok, G.S. (2005). Proceedings of the 2nd NIST LADAR Performance Evaluation Workshop – March 15-16, 2005. Retrieved March 17, 2009 from http://www.bfrl.nist.gov/861/CMAG/LADAR_workshop/NISTIR7266_final.pdf

Faber, P. & Fisher, R.B. A buyer's guide to Euclidean elliptical cylindrical and conical surface fitting. *Proc Br Mach Vis Conf BMVC01, Manchester, September* (2001), pp. 521–530

Feddema, J., & Little, C. (1997). Rapid World Modeling: Fitting Range Data to Geometric Primitives. *IEEE International Conference on Robotics and Automation, Albequerque, NM.*

Fitzgibbon, A., Pilou, M., & Fisher, R.B. (1999). Direct Least Square Fitting of Ellipses. Pattern Analysis and Machine Intelligence, Vol. 21, No. 5, May 1999.

Gao, C., Hoffman, I., Miller, T., Panzarella, T., & Spletzer, J. (2007). Performance Characterization of LIDAR Based Localization for Docking a Smart Wheelchair System. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007), Workshop on Assistive Technologies: Rehabilitation and Assistive Robotics*.

Gao, C., Hoffman, I., Miller, T., Panzarella, T., & Spletzer, J. (2008). Autonomous Docking of a Smart Wheelchair for the Automated Transport and Retrieval System (ATRS). *Journal of Field Robotics, Volume 25, No. 4-5, Pages 203-222*.

Gachter, S. (2005). Results on Range Image Segmentation for Service Robots. *Proceedings of the Fourth IEEE International Conference on Computer Vision Systems.* Retrieved on April 23, 2009, from http://asl.epfl.ch

Kasa, I. (1976). A circle fitting procedure and its error analysis. *IEEE Tranactions on Instrument Measurement*, *vol. IM-25, no. 3, pp. 8–14*.

Kato, H., & Billinghurst, M. (1999). Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System. *Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality '99, pp. 85-94, October 20-21*.

Lalonde, V,N., Huber, D.F., & Hebert, M. (2006) Natural Terrain Classification Using Three

Dimensional Ladar Data for Ground Robot Mobility. *Journal of Field Robotics 23(10) 839-861.*

Nguyen, V., Martinelli, A., Tomatis, N., & Siegwart, R. (2007). A Comparison of Line Extraction Algorithms using 2D Laser Rangefinder for Indoor Mobile Robotics. *Autonomous Robots. Vol.23, Issue 2 (August 2007). pp. 97-111.*

O'Leary, P., & Murray-Zsombor, P. (2004). Direct and Specific Least-Square Fitting of Hyperbolae and Ellipses. *Journal of Electronic Imaging 13(3), 492-503.*

Owen, C.B., Xiao, F., Middlin, P. (2009). What is the best fiducial? Retrieved on May 20, 2009 from http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=01107021

Premebida, C., & Nunes, U. (2005) Segmentation and Geometric Primtives Extraction from 2D Laser Range Data for Mobile Robot Applications. *Robotica 2005 – Actas do Encontro Cientifico.* Retrieved on 15 May 2009 from http://cyberc3.sjtu.edu.cn/doc/paper/ISR-Robotica05.pdf

Rusu, C., Tico, M., Kuosmanen, P., & Delp, E. (2003). Classical geometrical Approach to Circle Fitting - Review and New Developments. *Journal of Electronic Imaging 12(1), 179-193.*

Taubin, G. (1991). Estimation of Planar Curves, Surfaces, and Nonplanar Space Curves Defined by Implicit Equations with Applications to Edge and Range Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13, No. 11, November 1991.*

Steinbis, J., William, H., & Vincent, T.L. (2008). 3D Fiducials for Scalable AR Visual Tracking. *IEEE International Symposium on Mixed and Augmented Reality 2008, 15-18 September, Cambridge, UK.*

Zhao, J. , Roberts P.T.E., Colchester, A.C.F., Holton-Tainter, K. (1996). Registration Updating Using Marker Pins in a Video Based Neurosurgical Guidance System (VISLAN). Lecture Notes in Computer Science; Vol. 1131. *Proceedings of the 4th International Conference on Visualization in Biomedical Computing* pp. 287-196.